

Amendments to the Claims:

Listing of Claims:

1. (Canceled)

2. (Canceled).

3. (Canceled).

4. (Canceled).

5. (Previously presented) A method for transforming and electronic document comprising the steps of:

providing a visual representation of an original electronic document to a user;

receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said one or more virtual tags and said one or more transformation rules are determined by the steps of:

a. selecting one or more document elements for inclusion or exclusion in said virtual page from said visual representation of said original electronic document using a graphical user interface;

b. identifying said selected document elements using features of a personal data content mining (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual page;

c. collecting said identified document elements into a set; and

d. applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags and generate said one or more transformation rules.

6. (Original) The method of claim 5 wherein after said step d. of applying a classification algorithm further comprising the steps of:

- e. indicating said one or more virtual tags to said user at said visual representation; and
- f. approving said indicated one or more virtual tags; or
- g. disapproving said indicated one or more virtual tags, wherein when said indicated one or more virtual tags are disapproved repeating step a. through step e.

7. (Original) The method of claim 5 wherein said original electronic document is an original Web page said PDCM feature set comprises element description space features.

8. (Original) The method of claim 7 wherein said element description space features comprise one or more of the following features:

bold, not bold, italic, not italic, underline, not underline, superscript, subscript, normal type, number of links encountered before which document element within a current nested structure, size of a font, foreground color, background color, font face, surrounding header level, immediately preceding header level, immediately preceding comment text, table body, header, footer, caption, not a caption, cascading style sheet class, beginning of the current nested structure, amount of preceding visual space, pattern of preceding visual breaks, number of preceding visual breaks, path through a nested structure of said original Web page, table row at a document structure depth, table column at a document structure depth, and item count at a document structure depth.

9. (Original) The method of claim 5 wherein said original electronic document is an original Web page said PDCM feature set comprises path feature space features.

10. (Original) The method of claim 9 wherein said path feature space features comprise one or more of the following features:

a sequence, number of line breaks in a sequence, number of table cells in one row in a sequence, number of table cells in one column in a sequence, relativized feature space attributes, and number of preceding visual breaks of an item list number at a document structure depth.

11. (Original) The method of claim 5 further comprising the steps of:
determining stability of each of said features of said PDCM feature set; and

selecting said features of said PDCM feature set having a highest stability in said step d. of applying said classification algorithm.

12. (Previously presented) A method for transforming and electronic document comprising the steps of:

providing a visual representation of an original electronic document to a user;

receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said original electronic document is an original Web page said one or more virtual tags and said one or more transformation rules are determined by:

determining structural relationships of said original Web page to form a tree structure;

selecting one or more structural objects from said visual presentation of said original Web page;

selecting one or more document elements for inclusion or exclusion in said virtual Web page from said visual representation of said original Web page using a graphical user interface;

identifying said selected document elements using features of personal data content mining a (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual Web page;

collecting said identified document elements into a set;

applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags as one or more first virtual tags;

determining one or more second virtual tags from said feedback and said one or more structural objects;

associating said one or more second virtual tags to said tree structure; and

applying learning to associate said one or more first virtual tags to said one or more second virtual tags and to generate said one or more transformation rules.

13. (Canceled).

14. (Canceled).

15. (Canceled).

16. (Previously presented) A method for transforming and electronic document comprising the steps of:

providing a visual representation of an original electronic document to a user;

receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said one or more virtual tags and said one or more transformation rules are determined by the steps of:

a. selecting one or more document elements for inclusion or exclusion in said virtual page from said visual representation of said original electronic document using a graphical user interface;

b. identifying said selected document elements using features of a personal data content mining (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual page;

c. collecting said identified document elements into a set; and

d. applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags and generate said one or more transformation rules wherein said one or more transformation rules are applied to a more recent version of said original Web page.

17. (Canceled).

18. (Canceled).

19. (Currently amended) ~~The method of claim 1~~ A method for transforming and a dynamically changing electronic document comprising the steps of:

providing a visual representation of an original one or more instances of a dynamically changing electronic document to a user;

receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said one or more instances of said electronic document;

constructing one or more transformation rules using said feedback, and said one or more virtual tags extraction rules defining transformation of said electronic document; and

applying said one or more extraction transformation rules to said one or more instances of said electronic document, a second electronic document having a similar structure as said one or more instances of said document or future instances versions of said original electronic document to generate a virtual page of customized content extracted from said one or more instances of said electronic document, said second electronic document having a similar structure as said original document or said future versions of said electronic document; and

providing a visual representation of said virtual page wherein said one or more virtual tags are generated by the steps of:

categorizing all elements of said one or more instances of said electronic document as a plurality of OLAP cubes;

determining assignment of said feedback to said OLAP cubes; and

browsing said cubes to create said one or more virtual tags.

20. (Canceled).

21. (Canceled).

22. (Canceled).

23. (Canceled).

24. (Canceled).

25. (Canceled).

26. (Canceled).

27. (Canceled).
28. (Canceled).
29. (Canceled).
30. (Canceled).
31. (Canceled).
32. (Canceled).
33. (Canceled).
34. (Canceled).
35. (Canceled).
36. (Canceled).
37. (Canceled).
38. (Previously presented) A system for transforming an electronic document comprising:
 - means for providing a visual representation of an original electronic document to a user;
 - means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;
 - means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and
 - means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said means for constructing one or more transformation rules comprises:
 - means for selecting one or more document elements for inclusion or exclusion in said virtual page from said visual representation of said original electronic document using a graphical user interface;
 - means for identifying said selected document elements using features of a personal data content mining (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual page;

means for collecting said identified document elements into a set; and

means for applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags and generate said one or more transformation rules.

39. (Previously presented) The system of claim 38 further comprising:

means for indicating said one or more virtual tags to said user at said visual representation; and

means for approving said indicated one or more virtual tags; or

means for disapproving said indicated one or more virtual tags, wherein when said indicated one or more virtual tags are disapproved determining revised one or more virtual tags and applying said determined revised one or more virtual tags to said means for constructing one or more transformation rules.

40. (Original) The system of claim 38 wherein said electronic document is a Web page PDCM feature set comprises element description space features.

41. (Original) The system of claim 40 wherein said element description space features comprise one or more of the following features:

bold, not bold, italic, not italic, underline, not underline, superscript, subscript, normal type, number of links encountered before which document element within a current nested structure, size of a font, foreground color, background color, font face, surrounding header level, immediately preceding header level, immediately preceding comment text, table body, header, footer, caption, not a caption, cascading style sheet class, beginning of the current nested structure, amount of preceding visual space, pattern of preceding visual breaks, number of preceding visual breaks, path through a nested structure of said original Web page, table row at a document structure depth, table column at a document structure depth, and item count at a document structure depth.

42. (Original) The system of claim 38 wherein said original electronic document is an original Web page said PDCM feature set comprises path feature space features.

43. (Original) The system of claim 42 wherein said path feature space features comprise one or more of the following features:

a sequence, number of line breaks in a sequence, number of table cells in one row in a sequence, number of table cells in one column in a sequence, relativized feature space attributes, and number of preceding visual breaks of an item list number at a document structure depth.

44. (Original) The system of claim 38 further comprising:

means for determining stability of each of said features of said PDCM feature set; and

means for selecting said features of said PDCM feature set having a highest stability in said means for applying said classification algorithm.

45. (Previously presented) A system for transforming an electronic document comprising:

means for providing a visual representation of an original electronic document to a user;

means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said original electronic document is an original Web page said one or more virtual tags and said one or more transformation rules are determined by:

means for determining structural relationships of said original Web page to form a tree structure;

means for selecting one or more structural objects from said visual presentation of said original Web page;

means for selecting one or more document elements for inclusion or exclusion in said virtual Web page from said visual representation of said original Web page using a graphical user interface;

means for identifying said selected document elements using features of personal data content mining a (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual Web page;

means for collecting said identified document elements into a set;

means for applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags as one or more first virtual tags;

means for determining one or more second virtual tags from said feedback and said one or more structural objects;

means for associating said one or more second virtual tags to said tree structure; and

means for applying learning to associate said one or more first virtual tags to said one or more second virtual tags and to generate said one or more transformation rules.

46. (Canceled).

47. (Canceled).

48. (Canceled).

49. (Previously presented) A system for transforming an electronic document comprising:

means for providing a visual representation of an original electronic document to a user;

means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said means for constructing one or more transformation,

a first said one or more virtual tags is a portion of said original electronic document to be cut and a second one of said one or more virtual tags is a portion of said electronic document to be pasted and said one or more transformation rules being constructed from said first virtual tag and said second virtual tag for determining a cut and paste operation, wherein said one or more transformation rules are applied to a more recent version of said original Web page.

50. (Canceled).

51. (Canceled).

52. (Currently amended) ~~The system of claim 34~~ A system for transforming an a dynamically changing electronic document comprising:

means for providing a visual representation of an original one or more instances of a dynamically changing electronic document to a user;

means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said one or more instances of said electronic document;

means for constructing one or more transformation rules using said feedback, and said one or more transformation rules defining transformation of said electronic document virtual tags; and

means for applying said one or more extraction transformation rules to said one or more instances of said electronic document, a second electronic document having a similar structure as said one or more instances of said electronic document or future instances versions of said original electronic document to generate a virtual page of customized content extracted from said one or more instances of said electronic document, said second electronic document having a similar structure as said original document or said future versions of said one or more instances of said electronic document wherein said one or more virtual tags are generated by:

means for categorizing all elements of said one or more instances of said electronic document as a plurality of OLAP cubes;

means for determining assignment of said feedback to said OLAP cubes; and

means for browsing said OLAP cubes to create said one or more virtual tags.

53. (Canceled).

54. (Canceled).

55. (Canceled).

56. (Canceled).

57. (Canceled).

58. (Canceled).

59. (Canceled).

60. (Canceled).

61. (Canceled).

62. (Canceled).

63. (Canceled).

64. (Canceled).

65. (Canceled).

66. (Canceled).

67. (Canceled).

68. (Canceled).

69. (Canceled).

70. (Canceled).

71. (Previously presented) A computer program product for transforming an electronic document comprising:

means for providing a visual representation of an original electronic document to a user;

means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said means for constructing one or more transformation,

wherein said means for constructing one or more transformation rules comprises:

means for selecting one or more document elements for inclusion or exclusion in said virtual page from said visual representation of said original electronic document using a graphical user interface;

means for identifying said selected document elements using features of a personal data content mining (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual page;

means for collecting said identified document elements into a set; and

means for applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags and generate said one or more transformation rules.

72. (Original) The computer program product of claim 71 further comprising:

means for indicating said one or more virtual tags to said user at said visual representation; and

means for approving said indicated one or more virtual tags; or

means for disapproving said indicated one or more virtual tags, wherein when said indicated one or more virtual tags are disapproved determining revised one or more virtual tags and applying said means for indicated said one or more virtual tags to said user at said visual representation and means for approving said indicated virtual tags.

73. (Original) The computer program product of claim 71 wherein said original electronic document is an original Web page said PDCM feature set comprises element description space features.

74. (Original) The computer program product of claim 73 wherein said element description space features comprise one or more of the following features:

bold, not bold, italic, not italic, underline, not underline, superscript, subscript, normal type, number of links encountered before which document element within a current nested structure, size of a font, foreground color, background color, font face, surrounding header level, immediately preceding header level, immediately preceding comment text, table body, header, footer, caption, not a caption, cascading style sheet class, beginning of the current nested structure, amount of preceding visual space, pattern of preceding visual breaks, number of preceding visual breaks, path through a nested structure of said original Web page, table row at a document structure depth, table column at a document structure depth, and item count at a document structure depth.

75. (Original) The computer program product of claim 71 wherein said original electronic document is an original Web page said PDCM feature set comprises path feature space features.

76. (Original) The computer program product of claim 75 wherein said path feature space features comprise one or more of the following features:

a sequence, number of line breaks in a sequence, number of table cells in one row in a sequence, number of table cells in one column in a sequence, relativized feature space attributes, and number of preceding visual breaks of an item list number at a document structure depth.

77. (Original) The computer program product of claim 71 further comprising:
means for determining stability of each of said features of said PDCM feature set; and
means for selecting said features of said PDCM feature set having a highest stability in said means for applying said classification algorithm.

78. (Previously presented) A computer program product for transforming an electronic document comprising:

means for providing a visual representation of an original electronic document to a user;
means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content wherein said means for constructing one or more transformation,

wherein said original electronic document is an original Web page said one or more virtual tags and said one or more transformation rules are determined by:

means for determining structural relationships of said original Web page to form a tree structure;

means for selecting one or more structural objects from said visual presentation of said original Web page;

means for selecting one or more document elements for inclusion or exclusion in said virtual Web page from said visual representation of said original Web page using a graphical user interface;

means for identifying said selected document elements using features of personal data content mining a (PDCM) feature set and an intent of said user to include or exclude said document element in said virtual Web page;

means for collecting said identified document elements into a set;

means for applying a classification algorithm to said set to classify said one or more document elements into a respective said one or more virtual tags as one or more first virtual tags;

means for determining one or more second virtual tags from said feedback and said one or more structural objects;

means for associating said one or more second virtual tags to said tree structure; and

means for applying learning to associate said one or more first virtual tags to said one or more second virtual tags and to generate said one or more transformation rules.

79. (Canceled).

80. (Canceled).

81. (Canceled).

82. (Previously presented) A computer program product for transforming an electronic document comprising:

means for providing a visual representation of an original electronic document to a user;

means for receiving feedback from interaction by said user with said visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said original electronic document;

means for constructing one or more transformation rules using said feedback, said one or more transformation rules defining transformation of said electronic document; and

means for applying said one or more transformation rules to said electronic document, a second electronic document or future instances of said original document to generate a virtual page of customized content;

a first said one or more virtual tags is a portion of said original electronic document to be cut and a second one of said one or more virtual tags is a portion of said electronic document to

be pasted and said one or more transformation rules being constructed from said first virtual tag and said second virtual tag for determining a cut and paste operation;

wherein said one or more transformation rules are applied to a more recent version of said original Web page.

83. (Canceled).

84. (Canceled).

85. (Currently amended) ~~The computer program product of claim 70~~ A computer program product for transforming a dynamically changing electronic document comprising:

means for providing a visual representation of an one or more instances of a dynamically changing electronic document to a user;

means for receiving feedback from interaction by the user with the visual representation, said feedback is used to generate one or more virtual tags, said virtual tags identifying features of a portion of said one or more instances of said electronic document;

means for constructing one or more transformation rules using said feedback and said one or more virtual tags; and

means for applying said one or more transformation rules to said one or more instances of said electronic document, a second electronic document having a similar structure as said one or more instances of said electronic document or future versions of said electronic document to generate a virtual page of customized content extracted from said one or more instances of said electronic document, said second electronic document having a similar structure as said one or more instances of said electronic document or said future versions of said electronic document;

means for storing said one or more virtual tags with said one or more transformation rules as a respective one or more virtual tag objects in a virtual repository; and

means for retrieving said one or more stored virtual tag objects from said virtual repository when subsequently accessing said electronic document, said stored one or more transformation rules being used to generate said virtual page wherein said one or more virtual tags are generated by:

means for categorizing all elements of said one or more instances of said electronic document as a plurality of OLAP cubes;

means for determining assignment of said feedback to said OLAP cubes; and

means for browsing said OLAP cubes to create said one or more virtual tags.

86. (Canceled).

87. (Canceled).

88. (Canceled).

89. (Canceled).

90. (Canceled).

91. (Canceled).

92. (Canceled).

93. (Canceled).

94. (Canceled).

95. (Canceled).

96. (Canceled).

97. (Canceled).

98. (Canceled).

99. (Canceled).

100. (Canceled).

101. (Canceled).